

OIPe

(2)

RAW SEQUENCE LISTING DATE: 02/08/2001
 PATENT APPLICATION: US/09/761,117 TIME: 17:26:46

Input Set : A:\Cpg.pto
 Output Set: N:\CRF3\02082001\I761117.raw

SEQUENCE LISTING

4 (1) GENERAL INFORMATION:
 6 (i) APPLICANT: Dalla-Favera, Riccardo
 7 Chaganti, Raju S.K.
 9 (ii) TITLE OF INVENTION: CLONING AND USES OF THE GENETIC LOCUS
 10 bcl-6
 12 (iii) NUMBER OF SEQUENCES: 9
 14 (iv) CORRESPONDENCE ADDRESS:
 15 (A) ADDRESSEE: Cooper & Dunham LLP
 16 (B) STREET: 1185 Avenue of the Americas
 17 (C) CITY: New York
 18 (D) STATE: New York
 19 (E) COUNTRY: United States of America
 20 (F) ZIP: 10036
 22 (v) COMPUTER READABLE FORM:
 23 (A) MEDIUM TYPE: Floppy disk
 24 (B) COMPUTER: IBM PC compatible
 25 (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 26 (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
 28 (vi) CURRENT APPLICATION DATA:
 C--> 29 (A) APPLICATION NUMBER: US/09/761,117
 C--> 30 (B) FILING DATE: 16-Jan-2001
 31 (C) CLASSIFICATION:
 33 (viii) ATTORNEY/AGENT INFORMATION:
 34 (A) NAME: White, John P.
 35 (B) REGISTRATION NUMBER: 28,678
 36 (C) REFERENCE/DOCKET NUMBER: 0575/43771-A-PCT-US-Y
 38 (ix) TELECOMMUNICATION INFORMATION:
 39 (A) TELEPHONE: (212) 278-0400
 40 (B) TELEFAX: (212) 391-0525
 41 (C) TELEX: 422523 COOP UI
 44 (2) INFORMATION FOR SEQ ID NO: 1:
 46 (i) SEQUENCE CHARACTERISTICS:
 47 (A) LENGTH: 3720 base pairs
 48 (B) TYPE: nucleic acid
 49 (C) STRANDEDNESS: double
 50 (D) TOPOLOGY: linear
 52 (ii) MOLECULE TYPE: cDNA
 54 (ix) FEATURE:
 55 (A) NAME/KEY: CDS
 56 (B) LOCATION: 328..2445
 58 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 60 GGCCCTCGA GCCTCGAACC GGAACCTCCA AATCCGAGAC GCTCTGCTTA TGAGGACCTC 60
 62 GAAATATGCC GGCCAGTGAA AAAATCTTAT GGCTTTGAGG GCTTTGGTT GGCCAGGGGC 120
 64 AGTAAAAATC TCGGAGAGCT GACACCAAGT CCTCCCTGC CACGTAGCAG TGGTAAAGTC 180
 66 CGAAGCTCAA ATTCCGAGAA TTGAGCTCTG TTGATTCTTA GAACTGGGGT TCTTAGAAGT 240
 68 GGTGATGCAA GAAGTTTCTA GGAAAGGCCG GACACCAGGT TTTGAGCAAA ATTTTGACT 300

RAW SEQUENCE LISTING

DATE: 02/08/2001

PATENT APPLICATION: US/09/761,117

TIME: 17:26:46

Input Set : A:\Cpg.pto

Output Set: N:\CRF3\02082001\I761117.raw

```

70 GTGAAGCAAG GCATTGGTGA AGACAAA \ ATG GCC TCG CCG GCT GAC AGC TGT      351
71                                     Met Ala Ser Pro Ala Asp Ser Cys
72                                     1                               5
74 ATC CAG TTC ACC CGC CAT GCC AGG GAT GTT CTT CTC AAC CTT AAT CGT      399
75 Ile Gln Phe Thr Arg His Ala Arg Asp Val Leu Leu Asn Leu Asn Arg
76      10                      15                      20
78 CTC CGG AGT CGA GAC ATC TTG ACT GAT GTT GTC ATT GTT GTG AGC CGT      447
79 Leu Arg Ser Arg Asp Ile Leu Thr Asp Val Val Ile Val Val Ser Arg
80  25                      30                      35                      40
82 GAG CAG TTT AGA GCC CAT AAA ACG GTC CTC ATG GCC TGG AGA GGC CTG      495
83 Glu Gln Phe Arg Ala His Lys Thr Val Leu Met Ala Trp Arg Gly Leu
84      45                      50                      55
86 TTC TAT AGC ATC TTT ACA GAC CAG TTG AAA TGC AAC CTT AGT GTG ATC      543
87 Phe Tyr Ser Ile Phe Thr Asp Gln Leu Lys Cys Asn Leu Ser Val Ile
88      60                      65                      70
90 AAT CTA GAT CCT GAG ATC AAC CCT GAG GGA TTC TGC ATC CTC CTG GAC      591
91 Asn Leu Asp Pro Glu Ile Asn Pro Glu Gly Phe Cys Ile Leu Leu Asp
92      75                      80                      85
94 TTC ATG TAC ACA TCT CGG CTC AAT TTG CGG GAG GGC AAC ATC ATG GCT      639
95 Phe Met Tyr Thr Ser Arg Leu Asn Leu Arg Glu Gly Asn Ile Met Ala
96      90                      95                      100
98 GTG ATG GCC ACG GCT ATG TAC CTG CAG ATG GAG CAT GTT GTG GAC ACT      687
99 Val Met Ala Thr Ala Met Tyr Leu Gln Met Glu His Val Val Asp Thr
100 105                      110                      115                      120
102 TGC CGG AAG TTT ATT AAG GCC AGT GAA GCA GAG ATG GTT TCT GCC ATC      735
103 Cys Arg Lys Phe Ile Lys Ala Ser Glu Ala Glu Met Val Ser Ala Ile
104      125                      130                      135
106 AAG CCT CCT CGT GAA GAG TTC CTC AAC AGC CGG ATG CTG ATG CCC CAA      783
107 Lys Pro Pro Arg Glu Glu Phe Leu Asn Ser Arg Met Leu Met Pro Gln
108      140                      145                      150
110 GAC ATC ATG GCC TAT CGG GGT CGT GAG GTG GTG GAG AAC AAC CTG CCA      831
111 Asp Ile Met Ala Tyr Arg Gly Arg Glu Val Val Glu Asn Asn Leu Pro
112      155                      160                      165
114 CTG AGG AGC GCC CCT GGG TGT GAG AGC AGA GCC TTT GCC CCC AGC CTG      879
115 Leu Arg Ser Ala Pro Gly Cys Glu Ser Arg Ala Phe Ala Pro Ser Leu
116      170                      175                      180
118 TAC AGT GGC CTG TCC ACA CCG CCA GCC TCT TAT TCC ATG TAC AGC CAC      927
119 Tyr Ser Gly Leu Ser Thr Pro Pro Ala Ser Tyr Ser Met Tyr Ser His
120 185                      190                      195                      200
122 CTC CCT GTC AGC AGC CTC CTC TTC TCC GAT GAG GAG TTT CGG GAT GTC      975
123 Leu Pro Val Ser Ser Leu Leu Phe Ser Asp Glu Glu Phe Arg Asp Val
124      205                      210                      215
126 CGG ATG CCT GTG GCC AAC CCC TTC CCC AAG GAG CGG GCA CTC CCA TGT      1023
127 Arg Met Pro Val Ala Asn Pro Phe Pro Lys Glu Arg Ala Leu Pro Cys
128      220                      225                      230
130 GAT AGT GCC AGG CCA GTC CCT GGT GAG TAC AGC CGG CCG ACT TTG GAG      1071
131 Asp Ser Ala Arg Pro Val Pro Gly Glu Tyr Ser Arg Pro Thr Leu Glu
132      235                      240                      245
134 GTG TCC CCC AAT GTG TGC CAC AGC AAT ATC TAT TCA CCC AAG GAA ACA      1119

```

RAW SEQUENCE LISTING

DATE: 02/08/2001

PATENT APPLICATION: US/09/761,117

TIME: 17:26:46

Input Set : A:\Cpg.pto

Output Set: N:\CRF3\02082001\I761117.raw

135	Val	Ser	Pro	Asn	Val	Cys	His	Ser	Asn	Ile	Tyr	Ser	Pro	Lys	Glu	Thr	
136		250					255					260					
138	ATC	CCA	GAA	GAG	GCA	CGA	AGT	GAT	ATG	CAC	TAC	AGT	GTG	GCT	GAG	GGC	1167
139	Ile	Pro	Glu	Glu	Ala	Arg	Ser	Asp	Met	His	Tyr	Ser	Val	Ala	Glu	Gly	
140	265					270					275					280	
142	CTC	AAA	CCT	GCT	GCC	CCC	TCA	GCC	CGA	AAT	GCC	CCC	TAC	TTC	CCT	TGT	1215
143	Leu	Lys	Pro	Ala	Ala	Pro	Ser	Ala	Arg	Asn	Ala	Pro	Tyr	Phe	Pro	Cys	
144				285						290					295		
146	GAC	AAG	GCC	AGC	AAA	GAA	GAG	AGA	CCC	TCC	TCG	GAA	GAT	GAG	ATT		1263
147	Asp	Lys	Ala	Ser	Lys	Glu	Glu	Glu	Arg	Pro	Ser	Ser	Glu	Asp	Glu	Ile	
148			300						305					310			
150	GCC	CTG	CAT	TTC	GAG	CCC	CCC	AAT	GCA	CCC	CTG	AAC	CGG	AAG	GGT	CTG	1311
151	Ala	Leu	His	Phe	Glu	Pro	Pro	Asn	Ala	Pro	Leu	Asn	Arg	Lys	Gly	Leu	
152			315					320					325				
154	GTT	AGT	CCA	CAG	AGC	CCC	CAG	AAA	TCT	GAC	TGC	CAG	CCC	AAC	TCG	CCC	1359
155	Val	Ser	Pro	Gln	Ser	Pro	Gln	Lys	Ser	Asp	Cys	Gln	Pro	Asn	Ser	Pro	
156		330					335					340					
158	ACA	GAG	GCC	TGC	AGC	AGT	AAG	AAT	GCC	TGC	ATC	CTC	CAG	GGT	TCT	GGC	1407
159	Thr	Glu	Ala	Cys	Ser	Ser	Lys	Asn	Ala	Cys	Ile	Leu	Gln	Gly	Ser	Gly	
160	345				350					355					360		
162	TCC	CCT	CCA	GCC	AAG	AGC	CCC	ACT	GAC	CCC	AAA	GCC	TGC	AGC	TGG	AAG	1455
163	Ser	Pro	Pro	Ala	Lys	Ser	Pro	Thr	Asp	Pro	Lys	Ala	Cys	Ser	Trp	Lys	
164				365					370					375			
166	AAA	TAC	AAG	TTC	ATC	GTG	CTC	AAC	AGC	CTC	AAC	CAG	AAT	GCC	AAA	CCA	1503
167	Lys	Tyr	Lys	Phe	Ile	Val	Leu	Asn	Ser	Leu	Asn	Gln	Asn	Ala	Lys	Pro	
168			380						385					390			
170	GGG	GGG	CCT	GAG	CAG	GCT	GAG	CTG	GGC	CGC	CTT	TCC	CCA	CGA	GCC	TAC	1551
171	Gly	Gly	Pro	Glu	Gln	Ala	Glu	Gly	Arg	Leu	Ser	Pro	Arg	Ala	Tyr		
172			395					400					405				
174	ACG	GCC	CCA	CCT	GCC	TGC	CAG	CCA	CCC	ATG	GAG	CCT	GAG	AAC	CTT	GAC	1599
175	Thr	Ala	Pro	Pro	Ala	Cys	Gln	Pro	Pro	Met	Glu	Pro	Glu	Asn	Leu	Asp	
176		410					415					420					
178	CTC	CAG	TCC	CCA	ACC	AAG	CTG	AGT	GCC	AGC	GGG	GAG	GAC	TCC	ACC	ATC	1647
179	Leu	Gln	Ser	Pro	Thr	Lys	Leu	Ser	Ala	Ser	Gly	Glu	Asp	Ser	Thr	Ile	
180	425				430					435					440		
182	CCA	CAA	GCC	AGC	CGG	CTC	AAT	AAC	ATC	GTT	AAC	AGG	TCC	ATG	ACG	GGC	1695
183	Pro	Gln	Ala	Ser	Arg	Leu	Asn	Asn	Ile	Val	Asn	Arg	Ser	Met	Thr	Gly	
184				445					450					455			
186	TCT	CCC	CGC	AGC	AGC	AGC	GAG	AGC	CAC	TCA	CCA	CTC	TAC	ATG	CAC	CCC	1743
187	Ser	Pro	Arg	Ser	Ser	Ser	Glu	Ser	His	Ser	Pro	Leu	Tyr	Met	His	Pro	
188			460						465					470			
190	CCG	AAG	TGC	ACG	TCC	TGC	GGC	TCT	CAG	TCC	CCA	CAG	CAT	GCA	GAG	ATG	1791
191	Pro	Lys	Cys	Thr	Ser	Cys	Gly	Ser	Gln	Ser	Pro	Gln	His	Ala	Glu	Met	
192			475					480					485				
194	TGC	CTC	CAC	ACC	GCT	GGC	CCC	ACG	TTC	GCT	GAG	GAG	ATG	GGA	GAG	ACC	1839
195	Cys	Leu	His	Thr	Ala	Gly	Pro	Thr	Phe	Ala	Glu	Glu	Met	Gly	Glu	Thr	
196			490					495					500				
198	CAG	TCT	GAG	TAC	TCA	GAT	TCT	AGC	TGT	GAG	AAC	GGG	GCC	TTC	TTC	TGC	1887
199	Gln	Ser	Glu	Tyr	Ser	Asp	Ser	Ser	Cys	Glu	Asn	Gly	Ala	Phe	Phe	Cys	

RAW SEQUENCE LISTING

DATE: 02/08/2001

PATENT APPLICATION: US/09/761,117

TIME: 17:26:46

Input Set : A:\Cpg.pto

Output Set: N:\CRF3\02082001\I761117.raw

200	505		510		515		520		
202	AAT GAG TGT GAC TGC CGC TTC TCT GAG GAG GCC TCA CTC AAG AGG CAC							1935	
203	Asn Glu Cys Asp Cys Arg Phe Ser Glu Glu Ala Ser Leu Lys Arg His								
204		525		530		535			
206	ACG CTG CAG ACC CAC AGT GAC AAA CCC TAC AAG TGT GAC CGC TGC CAG							1983	
207	Thr Leu Gln Thr His Ser Asp Lys Pro Tyr Lys Cys Asp Arg Cys Gln								
208		540		545		550			
210	GCC TCC TTC CGC TAC AAG GGC AAC CTC GCC AGC CAC AAG ACC GTC CAT							2031	
211	Ala Ser Phe Arg Tyr Lys Gly Asn Leu Ala Ser His Lys Thr Val His								
212		555		560		565			
214	ACC GGT GAG AAA CCC TAT CGT TGC AAC ATC TGT GGG GCC CAG TTC AAC							2079	
215	Thr Gly Glu Lys Pro Tyr Arg Cys Asn Ile Cys Gly Ala Gln Phe Asn								
216		570		575		580			
218	CGG CCA GCC AAC CTG AAA ACC CAC ACT CGA ATT CAC TCT GGA GAG AAG							2127	
219	Arg Pro Ala Asn Leu Lys Thr His Thr Arg Ile His Ser Gly Glu Lys								
220	585		590		595		600		
222	CCC TAC AAA TGC GAA ACC TGC GGA GCC AGA TTT GTA CAG GTG GCC CAC							2175	
223	Pro Tyr Lys Cys Glu Thr Cys Gly Ala Arg Phe Val Gln Val Ala His								
224		605		610		615			
226	CTC CGT GCC CAT GTG CTT ATC CAC ACT GGT GAG AAG CCC TAT CCC TGT							2223	
227	Leu Arg Ala His Val Leu Ile His Thr Gly Glu Lys Pro Tyr Pro Cys								
228		620		625		630			
230	GAA ATC TGT GGC ACC CGT TTC CGG CAC CTT CAG ACT CTG AAG AGC CAC							2271	
231	Glu Ile Cys Gly Thr Arg Phe Arg His Leu Gln Thr Leu Lys Ser His								
232		635		640		645			
234	CTG CGA ATC CAC ACA GGA GAG AAA CCT TAC CAT TGT GAG AAG TGT AAC							2319	
235	Leu Arg Ile His Thr Gly Glu Lys Pro Tyr His Cys Glu Lys Cys Asn								
236		650		655		660			
238	CTG CAT TTC CGT CAC AAA AGC CAG CTG CGA CTT CAC TTG CGC CAG AAG							2367	
239	Leu His Phe Arg His Lys Ser Gln Leu Arg Leu His Leu Arg Gln Lys								
240	665		670		675		680		
242	CAT GGC GCC ATC ACC AAC ACC AAG GTG CAA TAC CGC GTG TCA GCC ACT							2415	
243	His Gly Ala Ile Thr Asn Thr Lys Val Gln Tyr Arg Val Ser Ala Thr								
244		685		690		695			
246	GAC CTG CCT CCG GAG CTC CCC AAA GCC TGC TGAAGCATGG AGTGTGATG							2465	
247	Asp Leu Pro Pro Glu Leu Pro Lys Ala Cys								
248		700		705					
250	CTTTCGTCTC CAGCCCCTTC TCAGAATCTA CCCAAAGGAT ACTGTAACAC TTTACAATGT							2525	
252	TCATCCCATG ATGTAGTGCC TCTTTCATCC ACTAGTGCAA ATCATAGCTG GGGGTGGGG							2585	
254	GTGGTGGGG TCGGGGCTG GGGGACTGGG AGCCGCAGCA GCTCCCCCTC CCCCCTGCC							2645	
256	ATAAAACATT AAGAAAATCA TATGCTTCT TCTCCTATGT GNNNNNNNNN NNNNNNNNNN							2705	
258	NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN							2765	
260	NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN							2825	
262	NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN							2885	
264	NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN							2945	
266	NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN							3005	
268	NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN							3065	
270	NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN							3125	
272	NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN							3185	

RAW SEQUENCE LISTING

DATE: 02/08/2001

PATENT APPLICATION: US/09/761,117

TIME: 17:26:46

Input Set : A:\Cpg.pto

Output Set: N:\CRF3\02082001\I761117.raw

```

274 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN 3245
276 NTTTAAGTAT TGCATCTGTA TAAGTAAGAA AATATTTTGT CTAAATGCC TCAGTGTATT 3305
278 TGTATTTTTT TGCAAGTGGG GGGTTACAAT TTACCCAGTG TGTATTAAAA AAAACCCAAA 3365
280 GAACCCAAAA ATCTCCAGAA GGAAAAATGT GTAATTTTGT TCTAGTTTTC AGTTTGTATA 3425
282 TACCCGTACA ACGTGTCTC ACGGTGCCTT TTTTCACGGA AGTTTTCAT GATGGGCGAG 3485
284 CGTGACCAT CCCTTTTGA AGTGTAGGCA GACACAGGA CTGAAGTTG TTAATACTA 3545
286 AACTCTCTTT GGGAATGTTT GTCTCATCCC ANTCTGCGTC ATGCTTGTGT GATACTACT 3605
288 CCGGAGACAG GGTTCGGCTG TGTCTAACT GCATTACGC GTTGTAATAA ATAGCTGTAC 3665
290 CAATATAAGA ATAAATGTT GGAAAGTCGC AAAAAAAAAA AAAAAA 3720
293 (2) INFORMATION FOR SEQ ID NO: 2:
295 (i) SEQUENCE CHARACTERISTICS:
296 (A) LENGTH: 706 amino acids
297 (B) TYPE: amino acid
298 (D) TOPOLOGY: linear
300 (ii) MOLECULE TYPE: protein
302 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
304 Met Ala Ser Pro Ala Asp Ser Cys Ile Gln Phe Thr Arg His Ala Arg
305 1 5 10 15
307 Asp Val Leu Leu Asn Leu Asn Arg Leu Arg Ser Arg Asp Ile Leu Thr
308 20 25 30
310 Asp Val Val Ile Val Val Ser Arg Glu Gln Phe Arg Ala His Lys Thr
311 35 40 45
313 Val Leu Met Ala Trp Arg Gly Leu Phe Tyr Ser Ile Phe Thr Asp Gln
314 50 55 60
316 Leu Lys Cys Asn Leu Ser Val Ile Asn Leu Asp Pro Glu Ile Asn Pro
317 65 70 75 80
319 Glu Gly Phe Cys Ile Leu Leu Asp Phe Met Tyr Thr Ser Arg Leu Asn
320 85 90 95
322 Leu Arg Glu Gly Asn Ile Met Ala Val Met Ala Thr Ala Met Tyr Leu
323 100 105 110
325 Gln Met Glu His Val Val Asp Thr Cys Arg Lys Phe Ile Lys Ala Ser
326 115 120 125
328 Glu Ala Glu Met Val Ser Ala Ile Lys Pro Pro Arg Glu Glu Phe Leu
329 130 135 140
331 Asn Ser Arg Met Leu Met Pro Gln Asp Ile Met Ala Tyr Arg Gly Arg
332 145 150 155 160
334 Glu Val Val Glu Asn Asn Leu Pro Leu Arg Ser Ala Pro Gly Cys Glu
335 165 170 175
338 Ser Arg Ala Phe Ala Pro Ser Leu Tyr Ser Gly Leu Ser Thr Pro Pro
339 180 185 190
341 Ala Ser Tyr Ser Met Tyr Ser His Leu Pro Val Ser Ser Leu Leu Phe
342 195 200 205
344 Ser Asp Glu Glu Phe Arg Asp Val Arg Met Pro Val Ala Asn Pro Phe
345 210 215 220
347 Pro Lys Glu Arg Ala Leu Pro Cys Asp Ser Ala Arg Pro Val Pro Gly
348 225 230 235 240
350 Glu Tyr Ser Arg Pro Thr Leu Glu Val Ser Pro Asn Val Cys His Ser
351 245 250 255
353 Asn Ile Tyr Ser Pro Lys Glu Thr Ile Pro Glu Glu Ala Arg Ser Asp

```

VERIFICATION SUMMARY DATE: 02/08/2001
PATENT APPLICATION: US/09/761,117 TIME: 17:26:47

Input Set : A:\Cpg.pto
Output Set: N:\CRF3\02082001\I761117.raw

L:29 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:30 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]